Bioestadística para no estadísticos
Tome decisiones con criterio

UNIVERSITAT POLITÈCNICA DE CATALUNYA
BARCELONATECH

CC BY NC SA

Capítulo 7

Inferencia y decisión

Erik Cobo, Jordi Cortés, José Antonio González y
Pilar Muñoz

Rosario Peláez, Marta Vilaró y Nerea Bielsa

Febrero 2015

Departament d'Estadística i Investigació Operativa
UNIVERSITAT POLITÈCNICA DE CATALUNYA

equator network

UPC

ELSEVIER DOYMA

MEDICINA CLINICA

TRIALS

Inferencia y decisión

Presentación

¿Qué información proporciona, a un clínico de Barcelona, los resultados obtenidos en un estudio previo realizado en Boston? La evolución de estos casos de Boston se puede conocer perfectamente, sin error. Pero esos casos ya han "evolucionado", no tiene interés predecir una evolución que ya ha sucedido. En cambio, sería muy interesante poder aplicar esos resultados pasados a unos nuevos casos. ¿Cómo hacerlo?

La inferencia estadística, para incorporar la información empírica, define los conceptos de muestra y población. Los valores obtenidos en una de las muchas posibles muestras permitirán estimar, con un cierto error cuantificable, el parámetro que caracteriza al conjunto de la población. La estadística pretende cuantificar, la información ("señal"), y el error ("ruido") que implica el proceso de generalización.

Este tema expone los conceptos fundamentales de inferencia estadística. Introduce la oscilación de los valores obtenidos en muestras aleatorias y cómo cuantificarla. Para ello, usando como ejemplo al estadístico más usual (la media muestral), estudia alrededor de qué valor oscila (su centro), cuánto varía (su dispersión) y qué forma adopta esta oscilación (su distribución).

El azar es clave, ya que permite estimadores sin sesgo y, quizás más importante, cuantificar la oscilación entre muestras.

Contribuciones: (1) la versión original de 2013 descansa en el libro de Bioestadística para No estadísticos de Elsevier de EC, JAG y PM, editada por JC y EC y revisada por RP; (2) la de enero de 2014 fue revisada por MV, JC y EC para incorporar mejoras y sugerencias anónimas; (3) la de septiembre de 2014 por NB y EC; y (4) la de febrero de 2015 por JC para incorporar mejoras de formato.

1. Introducción a la inferencia estadística

1.1. ¿Qué es la inferencia estadística?

Si, por ejemplo, se desea estimar el tiempo de crecimiento de un cierto tejido, se pueden utilizar dos procedimientos. El primero, teórico, consiste en *deducir* este tiempo a partir de los tiempos de división de sus células. El segundo procedimiento, empírico, consiste en *inducirlo* a partir de un número limitado de casos. Ahora bien, ¿hasta qué punto unas pocas pruebas permiten establecer modelos generales sobre el crecimiento de estos tejidos? O mejor, ¿cuánta información aportan? La inferencia estadística formaliza este proceso—que requiere: (1) definir, (2) cuantificar y (3) acotar sus riesgos.

Definición

La **inferencia** generaliza la información de una muestra a una población.

Historieta: Dos amigos caminan por el Pirineo y, al ver un caballo, uno de ellos comenta: “no sabía que los caballos de la Cerdaña fueran marrones y con las patas anchas”. Su amigo, que es lógico, le responde: “perdona, lo que no sabías es que en la Cerdaña hay, por lo menos, un caballo marrón de patas anchas”.

Lectura: Hasta hace relativamente poco, los filósofos lamentaban la falta de técnicas para saltar de las partes al todo. Para Hume, la inferencia era imposible; y para Russell, la inducción seguía siendo un problema de lógica no resuelto. A mediados del siglo pasado, Popper aportó un punto de vista algo más optimista: “sólo la refutación de una teoría puede ser inferida de datos empíricos y esta inferencia es puramente deductiva”. Hoy en día, en estudios bien diseñados, y ejecutados, la metodología estadística hace posible la inferencia.

1.2. Respuestas que ofrece la inferencia estadística

Veamos algunas preguntas que pueden ser contestadas con la ayuda de la metodología estadística. El ejemplo más sencillo estudiaría la distribución de una sola variable: ¿cuál es el valor de monóxido de carbono en el aire espirado por fumadores jóvenes? O bien, ¿cuál es la distribución de los valores de homocisteína plasmática en pacientes con lupus eritematoso?

Nota: si no existiera variabilidad, si la cantidad de monóxido de carbono espirado siempre fuera la misma, la inferencia sería inmediata: una observación bastaría para conocer el comportamiento de todas —aceptando la asunción de invarianza. La metodología estadística permite la diversidad, pero debe recurrir a otras premisas.

Ejemplo 1.1: No es necesario hacer un estudio estadístico para conocer la distribución de la variable “número de cerebros” que tiene cada uno de los habitantes de una ciudad.
Cada ciudadano tiene un cerebro y sólo uno. Así de fácil.

En cambio, sería terriblemente aburrido "decir toda la verdad" sobre la altura de una muestra de 23 pacientes: el primer caso mide 164 cm, el segundo, 173 cm; el tercero 168; ... y el vigésimo tercero, 192.

Ejercicio 1.1

Suponga que, en el Ejemplo 1.1, por no aburrir, decide hacer un resumen de los datos, ¿qué información le gustaría que este resumen le proporcionara: sobre el centro o sobre la dispersión?

Ejercicio 1.2

Proponga otro ejemplo en el que también sea conveniente hacer un resumen estadístico de los datos.

Conocer la distribución de una variable permitirá al clínico realizar de forma científica, por ejemplo, el diagnóstico, el tratamiento o el pronóstico.

Ejemplo 1.2: Si se conoce cuál es la distribución del tiempo de convalecencia tras cierta enfermedad, puede "adelantar" al paciente cuántos días tendrá sus facultades mermadas. Con la media, dirá al paciente cuál es su valor esperado. Y con la desviación típica, cuál es el error esperado: cuánto cabe esperar que un paciente típico se aleje de esa media.

1.3. Población, muestra e individuo

Las primeras definiciones son las de población, muestra y unidad.

Definiciones

Población: conjunto de todos los elementos, que cumplen ciertas propiedades comunes, entre los que se desea estudiar un determinado fenómeno.

Muestra: subconjunto de la población que es estudiado y a partir del cual se sacan conclusiones sobre las características de la población.

Unidad (individuo o caso): es cada uno de los elementos que componen la muestra y la población.

Población, muestra y unidad se contienen progresivamente, a la manera de las muñecas rusas. La población contiene la muestra y la muestra, las unidades. La diferencia es que, conceptualmente, hay un número ilimitado de muestras y de individuos. La población, sin embargo, es única y representa al conjunto que deseamos conocer.

Ejemplo 1.3: Costa et al. invitaron a participar en el estudio, de manera consecutiva, a todas las personas que acudieron al Centro de Extracciones del *Hospital Clínic i Universitari de Barcelona*, desde diferentes servicios, para la realización de una prueba de tolerancia oral a la glucosa (PTOG).
Cilla G et al.: "el estudio incluyó a mujeres que tuvieron un primer parto después de septiembre de 1989 y un segundo parto entre 2 y 8 años después en la Maternidad del Hospital Nuestra Señora de Aránzazu de San Sebastián (Guipúzcoa)".

Por su parte, las unidades no tienen por qué ser "individuos". Pueden ser hospitales, comarcas o visitas clínicas. Es muy importante definir con sumo cuidado estas unidades, ya que se podría llegar a conclusiones diferentes.

Ejemplo 1.4: Cierto facultativo presume de tener un razonable promedio de 7 pacientes por hora. Pero la asociación de usuarios ha preguntado a todos sus pacientes y ha obtenido un promedio de 9. ¡Y pudiera ser que todos digan la verdad, sin trampa!

Pongamos que este profesional tiene 3 horas de visita. En una de ellas ve a las primeras visitas, a razón de 3 por hora. En otra, recibe a las segundas visitas, 6 por hora. Y en la restante hora recibe las demás visitas, 12 por hora. Este facultativo ha definido como unidad del estudio la "hora de visita": el promedio de 3, 6 y 12 es, efectivamente, 7 pacientes por hora.

$$\sum_{i=1}^{3} \frac{X_i}{n} = \frac{3+6+12}{3} = 7 \; donde \; \sum_{i=1}^{3} indica\ la\ suma\ de\ los\ casos\ 1\ a\ 3$$

Los usuarios en cambio, han definido como unidad a cada uno de ellos, de forma que, en lugar de estudiar las 3 horas (unidades para el médico), estudian los 21 pacientes visitados por el médico. En sus 21 respuestas obtienen "3" en 3 pacientes; "6", en 6; y "12", en 12. Y el promedio es, efectivamente, 9 pacientes:

$$\sum_{i=1}^{21} \frac{X_i}{n} = \frac{3+3+3+6+6+6+6+6+6+12+\cdots+12}{3} = 9$$

Posiblemente la primera definición represente mejor la pregunta del clínico (¿qué promedio de pacientes visito yo por hora?); y la segunda, la del usuario (¿cuánto suele durar mi visita?). Ambas definiciones son correctas y válidas. Pero no son intercambiables y en cada estudio debe estar muy clara cuál es la unidad. Así, diferentes objetivos requieren diferentes cálculos, todos ellos lícitos y correctos, pero que no deben confundirse: siempre debe quedar bien clara la unidad del estudio.

Ejercicio 1.3

El colegio de odontólogos ha realizado un estudio aleatorio entre los pacientes de sus consultas en la semana anterior. De 1000 fichas analizadas, 500 habían tenido una visita el año anterior, por lo que concluyen que un 50% de la población acude al dentista cada año. ¿Qué opina? ¿Se puede conocer la frecuencia de visitas al odontólogo en la población general a partir de una muestra obtenida en las consultas?

Ejercicio 1.4

Los centros sanitarios de la Seguridad Social suelen realizar una encuesta de satisfacción a sus usuarios, cuyos resultados suelen ser altamente positivos. ¿Qué llevó al defensor del pueblo a realizar una encuesta en la población general? (Pista: defina la unidad de ambos estudios y medite sus diferencias).

Ejercicio 1.5

Para estimar la infección nosocomial, puede hacerse un estudio seleccionando algunos de los pacientes que ingresan o bien seleccionando algunas de las camas ocupadas en el hospital. ¿Cuáles son las unidades de ambos tipos de estudios? Asumiendo que los pacientes que están ingresados más tiempo tienen mayor tasa diaria de desarrollar la infección, ¿cuál de los dos estudios dará cifras más altas?

Recuerde

Dos estudios, para ser comparables, requieren la misma unidad.

1.4. Estadístico, estimador y parámetro

En el estudio de la información disponible, la inferencia estadística afronta el reto de abarcar un "universo" más amplio que los pocos casos de que dispone.

Definición

El indicador que se obtendría de cada posible muestra, se llama **estadístico**.

El indicador de la población que se desea conocer, se llama **parámetro**.

Recuerde

Parámetro refiere a la población; estadístico, a la muestra.

Por ejemplo, el término **esperanza** representa el "parámetro" que indica el centro de gravedad de una distribución poblacional;, pero el "estadístico" media refiere al **promedio** calculado en una muestra.

Nota: En ocasiones, la esperanza también recibe el nombre de media poblacional.

Ejemplo 1.4 (continuación): Supóngase que la probabilidad (π) de que un paciente con anticuerpos del SIDA tarde, en ciertas condiciones, más de 2 años en desarrollar los primeros síntomas es 0.50. Es decir, expresado en porcentajes, del 50%. Esta probabilidad es un **parámetro** que resume las expectativas del paciente y que representa una característica intrínseca de la enfermedad. Por otro lado, en una muestra de 25 pacientes de esa población, 15 han superado los 2 años. Este resultado de 15/25 = 0.60 (60%) representa la proporción, que es el **estadístico** o valor observado en la muestra.

Ejercicio 1.6

Proponer un ejemplo similar con media y esperanza.

El reto de la inferencia estadística es conocer el parámetro, que caracteriza al total de la población, a partir de estadísticos obtenidos en una muestra.

Nota: Si Vd. dispone de los datos de toda la población, es decir, si las conclusiones de su estudio aplican únicamente a estos casos y no desea aplicarlas a otros diferentes, Vd. no necesita saber lo que es la inferencia estadística. Pero vigile al hablar: no podrá establecer ninguna ley 'universal' que vaya más allá de sus propios datos.

Definición

Cuando un estadístico de una muestra se usa para conocer el valor de un parámetro de la población, recibe el nombre de **estimador**.

Nota: Cada muestra es fugaz, en el sentido de ser irrepetible y, en el fondo, irrelevante en sí misma. Una vez terminado el seguimiento de los pacientes de la muestra y cumplidas las responsabilidades sanitarias con ellos, el interés científico se centrará en conocer qué dicen estos casos sobre los pacientes futuros.

Ejemplo 1.5: Las encuestas electorales, a partir de unos pocos miles de entrevistados intentan conocer la tendencia de unos cuantos millones: el auténtico interés está en lo que votará toda la población. La importancia que tienen los pocos entrevistados es su capacidad para informar sobre la distribución poblacional de esta variable.

Recuerde

Se puede acceder al estadístico observado en la muestra; pero el auténtico objetivo, el parámetro de la población, no suele ser accesible.

Definición

La **inferencia estadística** cuantifica la información empírica (evidencia, o pruebas) que el estimador proporciona del parámetro.

Es tan importante distinguir si se trata de valores muestrales o poblacionales que se les dará diferente símbolo en un caso o en otro. (Tabla 1.1).

	Parámetro (θ) **(Población)**	**Estadístico ($\hat{\theta}$)** **(Muestra)**
Media	μ=E(X) esperanza	$\bar{X}$ media o promedio
Desviación típica	σ	S
Probabilidad	π probabilidad	P proporción

Tabla 1.1 Notación utilizada para distinguir parámetro y estadístico

Así, una proporción observada en una muestra informa sobre la probabilidad de la población. Pero, ¿cuánta información aporta? ¿El valor poblacional se acerca mucho o poco al valor del estadístico observado? La teoría de probabilidad permite cuantificar la información que un estadístico (proporción *P*) aporta sobre el valor desconocido del parámetro (probabilidad π), auténtico objetivo del estudio.

Si se sabe cuánto oscila un estadístico de una muestra a otra, se podrá cuantificar la información que contiene. Veamos la distribución del promedio obtenido en una muestra aleatoria simple (MAS).

1.5. Muestra aleatoria simple

Definición

Muestra Aleatoria Simple (MAS) es aquella en la que (1) todos los elementos de la población tienen la misma probabilidad de pertenecer; y (2) cualquier combinación de n elementos tiene la misma probabilidad de pertenecer a ella.

Nota: Todos los elementos de la muestra tienen la misma distribución, ya que vienen de la misma población.

Ejemplo 1.6: Imaginemos: 1) la población infinita de todos los posibles pacientes de una enfermedad; 2) un procedimiento aleatorio que selecciona de forma independiente n=1000 pacientes de esta población.

Contra-Ejemplo 1.7: Una asociación profesional con 25000 afiliados decide hacer un estudio para conocer qué proporción de ellos han recibido malos tratos en su trabajo. Diseña una muestra aleatoria de 2000 a los que les envía un cuestionario, que contestan sólo 500. Se puede saber que los 2000 representan a los 25000, pero se desconoce a quién representan estos 500 y, por tanto, qué información aportan sobre el total de la población.

Nota: La definición de la población a la que se desea aplicar los resultados puede cambiar la consideración de la muestra.

Contra-Ejemplo 1.8: Supongamos: 1) la población finita de los 80 pacientes de una enfermedad determinada de un centro hospitalario; 2) un proceso aleatorio de selección de 20 pacientes diferentes. Nótese que, al ser un muestreo sin reemplazamiento, al eliminar un paciente cada vez, la población de los pacientes susceptibles de ser seleccionados va variando, con lo que la variable no tiene la misma distribución para cada uno de los elementos de la muestra.

Ejemplo 1.9: En el fondo, el objetivo del estudio del Contra-Ejemplo 1.8 no puede ser conocer cómo se comportan estos 80 pacientes (tema vital para ellos y para el centro que los atiende, pero sin ningún interés para el resto de pacientes y centros). El objetivo del estudio debe ser más ambicioso, de manera que se puedan beneficiar los pacientes de otros profesionales y centros. Ahora, por un lado la situación se simplifica, ya que eliminar un elemento de esta población infinita prácticamente no modifica su distribución. Pero, por otro lado se complica, ya que debe tenerse en cuenta que los casos estudiados (sean 20 o sean 80) no son una muestra aleatoria de la población de todos los pacientes con la misma enfermedad. ¿Hasta qué punto los resultados son extrapolables?

> **Recuerde**
>
> Caso, muestra y población no se definen por separado, de forma aislada. Haga siempre la definición conjunta.

Volvamos a la definición de MAS. También resalta que la información aportada por las diferentes unidades deba ser independiente entre sí. Es decir, el valor obtenido en una observación no aporta información sobre el valor de otras observaciones. Este "no aportar información" debe entenderse como que la distribución de las restantes variables es la misma sea cual sea el valor observado.

Ejemplo 1.10: Sigamos con el ejemplo 1.6 de pacientes con una enfermedad. Cada uno de los elementos de la muestra aporta exactamente la misma información sobre la población: que cierto paciente tenga un valor elevado no implica que cualquier otro paciente lo deba tener ni más alto ni más bajo.

Contra-Ejemplo 1.11: En un estudio multi-céntrico, ¿puede creerse que el resultado de un paciente de un centro no aporta información sobre el resultado de otro paciente del mismo centro? O por el contrario, ¿es más razonable pensar que los resultados obtenidos en pacientes de un mismo centro son más similares que los de pacientes de centros diferentes? Si es éste último caso, la variable centro es una variable importante que debe ser tenida en cuenta en el análisis posterior.

Contra-Ejemplo 1.12: Preguntar la altura a 5 estudiantes proporciona 5 piezas de información que se van añadiendo. Pero preguntar otra vez a uno de estos no estudiantes no aportan información nueva, independiente de la ya disponible.

> **Recuerde**
>
> En una MAS: 1) las unidades se escogen al azar; 2) todas ellas tienen la misma probabilidad de ser escogidas; 3) todas las posibles combinaciones de elementos tienen la misma probabilidad de figurar en la muestra.

1.6. Inferencia estadística y proceso científico

La capacidad de la estadística para inferir formalmente desde unos pocos datos de la muestra a la totalidad de la población ha permitido un progreso espectacular en todas las ciencias. Hoy en día, se acepta como modelo de razonamiento científico el contenido en el siguiente esquema.

1. Descubrir el problema a investigar.
2. Documentar y definir el problema o hipótesis.
3. Deducir consecuencias contrastables de las hipótesis.
4. Diseñar la observación o la experimentación.
5. Recoger los datos.
6. Análisis de datos mediante inferencia estadística.
7. Interpretar.
8. Integrar en el cuerpo de conocimiento.

Cuadro 1.1 Pasos del método científico

Este modelo integra razonamiento inductivo y deductivo. El razonamiento deductivo es necesario, por ejemplo, para diseñar la recogida de datos. El inductivo, por su parte, es necesario para generalizar las observaciones obtenidas en unos cuantos elementos.

Lectura: Su versión profesional es el ciclo plan/do/check/act**:** (1 plan) tras un análisis de la situación de partida, el planificador establece una estrategia para alcanzar unos objetivos; (2 do) ejecuta el trabajo definido en el plan anterior; (3 check) evalúa el grado de obtención de objetivos; y (4 act) en base a los resultados del paso anterior elige una opción (quizás cambiar el plan por uno mejor o quizás aplicarlo tal cual en una escala mayor). Y vuelve a empezar el ciclo.

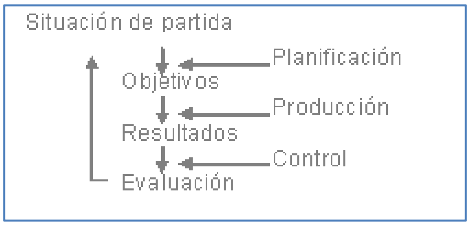

Figura 1.1 Esquema del método científico- técnico.

1.7. Posibles errores en la inferencia estadística

En Bioestadística, la definición previa de muestra aleatoria se enfrenta a varios retos. En primer lugar, los individuos tienen derecho a rechazar su participación en el estudio, o incluso a abandonarlo en cualquier momento. En segundo lugar, puede no haber una definición operativa de la población; por ejemplo: no hay ningún listado con todos los pacientes de una determinada enfermedad. Todos estos fenómenos no aleatorios pueden provocar distorsiones no aleatorias, llamadas **sesgos**.

Ejemplo 1.13: Si cierto número de casos no termina el estudio, el investigador debe dejarlo claro y analizar o, por lo menos discutir, hasta qué punto compromete las conclusiones.

Lectura: ¿Hasta qué punto debemos creernos las previsiones electorales que se publican en diferentes medios? A continuación, y respecto a las elecciones generales de octubre de 1989, figuran los resultados reales (parámetros poblacionales) junto a las previsiones (estimaciones basadas en muestras) publicadas por El Periódico de Catalunya y por La Vanguardia. El segundo diario, que se comprometía con un margen menos ambicioso (2%), cumplió. En cambio, el primero falló con dos formaciones: a pesar de que prometía un margen de error máximo de 1 punto, para el PP se distanció 6.7 puntos; y para IU, 1.2.

RESULTADOS (%)		**Previsiones**	
		EL PERIÓDICO 23/10/1989 n=9524 +2000 (%)	**LA VANGUARDIA** 23/10/1989 n=3262 (%)
PSOE	39.6	40.5	41.5
PP	25.8	19.1	25.0
CIU	5.0	4.9	4.5
IU	9.1	10.3	7.5
CDS	7.9	8.5	6.5
Margen		**± 1**	**± 2**

Tabla 1.2 Prospección de voto y resultados electorales de octubre de 1989

Las "fichas técnicas" de ambos estudios aportan explicaciones a estas diferencias: tipo de entrevista (personal o telefónica); método de selección de los casos; días en los que se realizó la encuesta; considerar la profesión como estrato; más encuestas en Cataluña donde el PP suele estar bajo;... Nótese que estas explicaciones se basan en argumentos sociológicos, no estadísticos. Un argumento estadístico es que, por pura mala suerte, la estimación de El Periódico, se alejó del auténtico valor.

La lectura anterior ilustra que en todo muestreo hay dos clases de errores: los debidos exclusivamente a las fluctuaciones del azar o **errores aleatorios;** y todos los demás, conocidos como **errores sistemáticos** o **sesgos**. La estadística ayuda a cuantificar la magnitud de los primeros. Controlar los segundos es una responsabilidad compartida entre la estadística y la disciplina aplicada. En el ejemplo anterior, la Sociología. En los estudios clínicos, el profesional sanitario debe razonar si las condiciones del estudio permiten negar la existencia de sesgos. La Figura 1.2 adelanta los posibles sesgos en un ensayo clínico —que se verán en el tema 10.

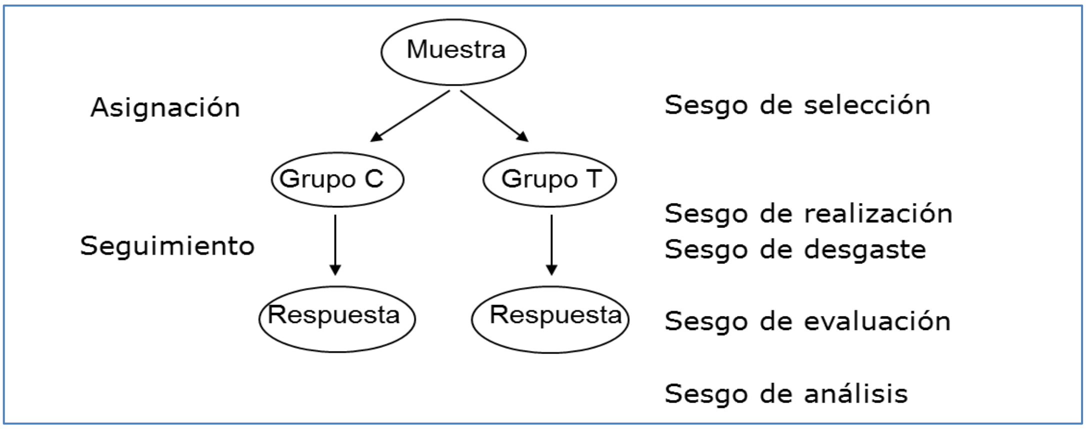

Figura 1.2 Algunos posibles sesgos

Y la **Tabla 1.3** algunos remedios.

Sesgo de …	selección	Asignación oculta al azar
	realización	Con placebo y enmascarado
	desgaste	Seguimiento completo
	evaluación	Enmascarada
	análisis	Plan de análisis previo a los datos

Tabla 1.3. Cómo evitar algunos tipos de sesgo

Recuerde

Al pasar de la muestra a la población, en el proceso inferencial hay dos posibles fuentes de errores: los *aleatorios* que la Estadística le ayudará a cuantificar; y los sistemáticos, o *sesgos*, cuya posible existencia debe Vd. estudiar a la luz de sus conocimientos clínicos.

1.8. Poblaciones implicadas en la inferencia estadística

Un estudio debe, en primer lugar, cuantificar la magnitud del error aleatorio; luego, justificar que esta magnitud es suficiente para los objetivos del estudio; y, en tercer lugar, defender la ausencia de sesgos. Cuando se dan estas condiciones, se dice que el estudio es válido. Ahora bien, ¿válido para qué conclusiones, las de los autores del estudio o las de aquellos que desean aplicarlo en una nueva población? Las siguientes definiciones de poblaciones, progresivamente más amplias, ayudan a clarificar estos conceptos.

Definiciones

Población **origen** de la muestra *("actual population")* o población muestrada es aquella población *imaginaria* de la que se hubiera obtenido, por extracción aleatoria 'pura', la muestra. Excluye, por ejemplo, a los casos que se niegan a participar en el estudio.

Población **objetivo** o diana *("target population")* es aquella población a la que los autores del estudian desean aplicar los resultados.

Poblaciones **externas** *("external populations")*: son otras poblaciones, quizás más amplias, a las que otros investigadores pueden desear aplicar los resultados.

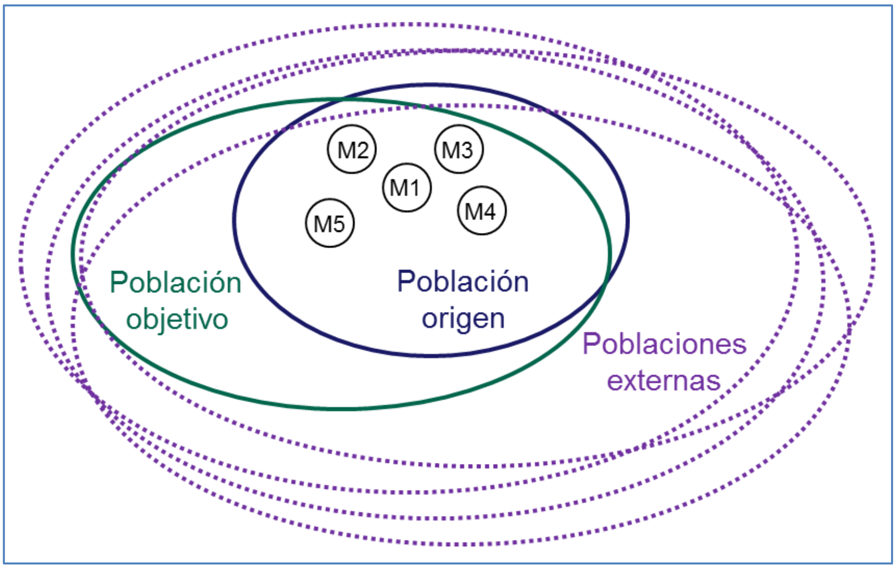

Figura 1.3 Esquema de las poblaciones y muestras (M1…M5) de un estudio

Como pueden no incluirse entre sí, no podemos representarlas con las muñecas rusas (Figura 1.3). Las posibles muestras SÍ están comprendidas en la población muestreada: si son aleatorias, la inferencia estadística permite este salto. Además el autor del estudio debe valorar posibles diferencias o sesgos entre las poblaciones origen y objetivo. Y los que deseen aplicar los resultados, con su propia población externa.

Ejemplo 1.14: Los ensayos clínicos suelen elegir casos entre 18 y 65 años. Así, los pacientes menores y mayores no forman parte de la población objetivo. Sí forman parte de la externa, los argumentos para aplicar a ellos los resultados del estudio son ajenos a la inferencia estadística.

Lectura: La definición de los criterios de selección es crucial: según Rafael Dal-Ré et al, "En general, la inclusión de enfermos en ensayos clínicos es un problema más importante de lo que los propios investigadores piensan, y siempre resulta más difícil de lo que en un principio se planeó."

Recuerde

Criterios de elegibilidad muy restrictivos dificultarán el reclutamiento y limitarán la aplicación posterior de los resultados.

Ejercicio 1.7

A partir de un artículo de investigación (p.e., "resultados a los 12 meses de un programa de deshabituación tabáquica en un centro de atención primaria") defina: una población externa, población objetivo, población muestreada y muestra.

Historieta: Si le preguntan "¿es representativa su muestra?", conteste con aplomo que sí: siempre hay una población origen para la que su muestra sería representativa. Las preguntas de interés son ¿representa a la población objetivo? ¿Cómo la ha definido? El reto es interpretar de forma clara y transparente a dónde le permiten llegar sus datos.

Definición

Error aleatorio es la variación entre los valores muestrales (estadísticos) obtenibles en las posibles muestras (centrados en el parámetro de la población origen).

Un **estudio es preciso** si el error aleatorio es pequeño.

Sesgo es una diferencia entre los valores del parámetro en las poblaciones origen y objetivo.

Hay **validez interna** si no tiene sesgos respecto a la población objetivo. Y hay **validez externa** si la ausencia de sesgos abarca a la población externa.

Ejercicio 1.8

Los textos médicos suelen estar basados en artículos científicos escritos desde centros de atención terciaria. ¿Puede este hecho provocar un sesgo?

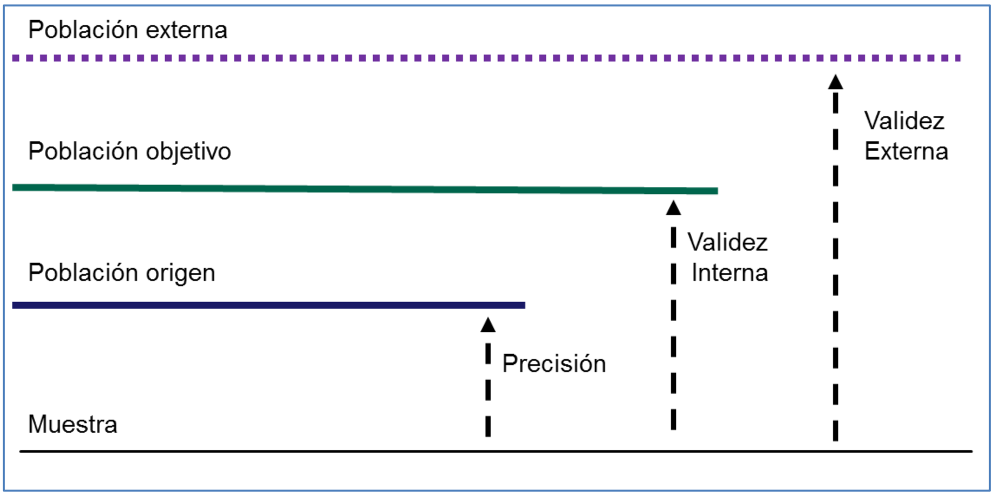

Figura 1.4 Relación entre propiedades de inferencia y poblaciones (Kleinbaum et al.)

Ejercicio 1.9

Autores, revisores y editores tienden a valorar más los estudios con resultados estadísticamente significativos. ¿Puede esta actitud provocar un sesgo?

Recuerde

La inferencia estadística sólo cuantifica la magnitud del error aleatorio.

Historieta: Seleccionar una muestra "al tuntún" no es lo mismo que hacerlo "al azar".

2. Estadístico media muestral

2.1. Distribución del estadístico media muestral $\bar{X}$

Recuerde

El cálculo del promedio o media muestral $\bar{X}$ es: $\bar{X} = \sum x_i / n$

Lectura: Mediavilla et al. La media del colesterol LDL (mmol/l) en los pacientes incluidos es $\bar{X}_{incl}$= 3.33 y la de los pacientes excluidos es $\bar{X}_{excl}$= 3.49.

Detengámonos un segundo: la media muestral, ¿tiene distribución? ¿Qué significa esto? La pregunta es si el "estadístico media muestral" es una constante o, por el contrario, se trata de una variable que debe ser caracterizada por su distribución. Como hay muchas muestras aleatorias posibles, cada una con diferentes casos, la media varía: es una variable.

Ejemplo 2.1: En el estudio anterior, si se obtienen dos muestras, las medias $\overline{X}$ de colesterol LDL serán algo diferentes, aunque se trate de casos de la misma población.

Ejemplo 2.2: Seleccionamos al azar n=100 pacientes con hipertensión y calculamos la media muestral o promedio de sus valores del colesterol LDL. Este valor será diferente que si obtenemos otra muestra de 100 pacientes y volvemos a calcular su media.

Historieta: Seleccionamos al azar un paciente y le preguntamos 100 veces por su edad (¡Pobre tipo! ¡Qué paciencia! ¿Y qué pensará de nosotros?) y hacemos el promedio. Ahora, cabe esperar que obtengamos el mismo valor si calculamos este promedio en otras 100 preguntas sobre su edad (se asume que el paciente es muy paciente, claro). Si en medio de tantas preguntas no ha cumplido años, no habrá variabilidad en la edad, y la media tampoco variará.

Ejemplo 2.3: Al paciente anterior, en lugar de preguntarle la edad, se le determina 5 veces la presión arterial. Ahora, los resultados podrían variar. Nótese que, en este caso, la variabilidad que se cuantifica hace referencia a la población de las diferentes mediciones en un mismo paciente, no a la población de pacientes. Es decir, cuantifica la variabilidad intra-paciente en lugar de la entre-paciente.

Nota: Esta variabilidad intra-paciente podría incluir cambios naturales del paciente, pero también error aleatorio de medida, diferente calibrado, distinto evaluador, etc. Dejamos su estudio para el curso de observacionales.

Así, las medias varían de una muestra a otra. Si se desea utilizar el estadístico promedio como estimador del parámetro poblacional "esperanza", esta variabilidad inducirá a errores lo que, por supuesto, nunca es deseable. Ahora bien, ¿se pueden cuantificar estos errores? O, lo que es más importante, ¿se puede limitar su magnitud? Para responder a estas dos preguntas cruciales, se debe contestar antes a otras más sencillas:

1) ¿Alrededor de qué valor varían? (Es decir, ¿cuál es su centro?)
2) ¿Varían mucho o poco alrededor de este valor? (Es decir, ¿cuál es su dispersión?)
3) ¿Qué forma tiene su distribución?

Nótese que, una vez aceptado que el estadístico promedio o media muestral $\overline{X}$ tiene una cierta distribución, las dos primeras preguntas se reducen a conocer su centro y dispersión poblacional. Veámoslas sucesivamente.

2.2. Centro de la distribución del estadístico $\bar{X}$

El centro poblacional del estadístico media muestral $\bar{X}$ recibe el nombre de *esperanza de* $\bar{X}$ y se representa por $E(\bar{X})$. Se sabe que, si la muestra es aleatoria simple, la esperanza de $\bar{X}$ coincide con la esperanza de X.

Fórmula

$E(\bar{X}) = E(X) = \mu$

Ejemplo 2.4: Si obtenemos MAS del colesterol LDL en 100 pacientes, el centro de las medias de todas las muestras ($E(\bar{X})$) coincide con la media poblacional del LDL ($E(X)=\mu$).

Recuerde

Si el muestreo es aleatorio, el centro de $\bar{X}$ coincide con el de *X*.

Ejercicio 2.1

¿Es deseable esta situación? ¿Qué utilidad puede tener este hecho?

En resumen, se sabe que el conjunto de las medias de todas las posibles muestras aleatorias tiene su centro, $E(\bar{X})$, en el mismo centro de la variable en estudio, $\mu=E(X)$. Así, cuando usamos la media de la muestra $\bar{X}$ para conocer la media de la población $\mu=E(X)$, hay el 'consuelo' de que el conjunto de todas las posibles muestras "apuntan" en la dirección correcta. Los errores serán tanto por exceso como por defecto. Pero, estos errores no tienen "favorito": hay equilibrio entre los positivos y los negativos.

2.3. $\bar{X}$ es un estimador insesgado de $\mu=E(X)$

Definición

Un **estimador** es **insesgado** si el centro de su distribución a lo largo de todas las posibles muestras coincide con el parámetro.

Ejemplo 2.5: Se desea estimar cómo evolucionan los ingresos de los médicos colegiados. Cada año, seleccionados al azar informan sobre su salario. El conjunto de las medias de todas las posibles muestras tiene como centro $E(\bar{X})$ al centro de los salarios, E(X).

Contra-Ejemplo 2.6: No sería correcto extrapolar estos resultados a otros colectivos con otros salarios. Si se hiciera, se cometería un sesgo igual a la diferencia entre los salarios medios de ambos colectivos.

Analogía: Sean dos lanzadores con arco que apuntan a sus respectivas dianas (Figura 2.1). El lanzador de la izquierda tiene un sesgo hacia la izquierda y arriba, mientras que el de la derecha está centrado.

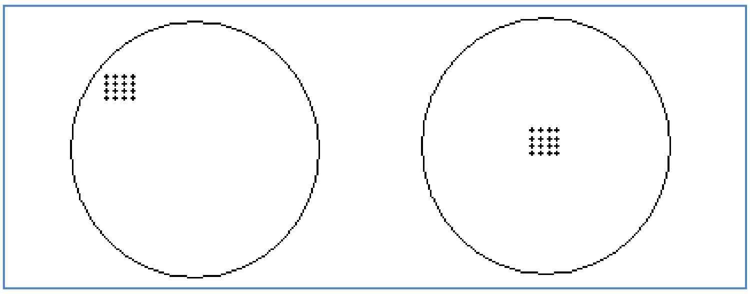

Figura 2.1 El arquero de la izquierda tiene mayor sesgo que el de la derecha.

Ejemplo 2.7: Dos informáticos han diseñado dos experimentos para conocer el tiempo de un nuevo algoritmo para decodificar el ADN. El primer informático analiza muestras del cromosoma 21, más corto. Mientras que el segundo selecciona muestras de todos los cromosomas. Las posibles muestras del primero tendrán medias muestrales, cuyo centro, $E(\bar{X})$, estará por debajo de la media poblacional, $\mu=E(X)$. Las del segundo informático estarán centradas en la auténtica media poblacional (**Figura 2.2**).

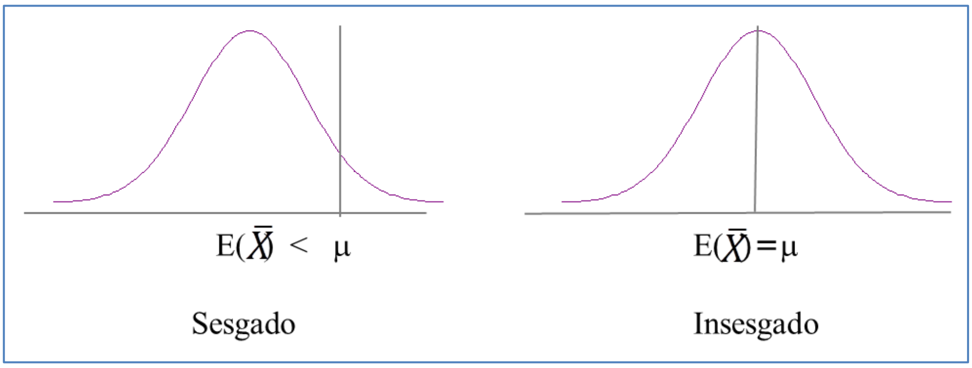

Figura 2.2 Estimador sesgado, $E(\bar{X}) \neq \mu$; e insesgado, $E(\bar{X}) = \mu$

Recuerde

Si la muestra es aleatoria, el promedio muestral $\bar{X}$ es insesgado.

Podría ser peor, claro, podría ser que las estimaciones apuntaran en dirección incorrecta. La ausencia de sesgo parece un pobre consuelo, ya que estas estimaciones no aciertan. Si no se puede garantizar que cada estimación acierte, ¿se puede por lo menos cuantificar la magnitud de su error? La varianza de $\bar{X}$ aportará esta información.

2.4. Dispersión de la distribución del estadístico $\bar{X}$

Ejercicio 2.2

¿Recuerda el cálculo de la varianza de una muestra?

REPASO: $S^2 = \sum(x_i - \bar{X})^2/(n-1)$

$S^2 = [\sum x_i^2 - (\sum x_i)^2/n]/(n-1)$ (más eficiente)

Practique ambas fórmulas, a mano y con R, para el ejemplo sencillo de n=5 alumnos que contestan que en su familia son 1, 2, 3, 4 y 5 hermanos. [Es una excepción, no se lo pedimos nunca. Hágalo. Convienen interiorirzarlas.]

¿Qué fórmula es más intuitiva? ¿Qué ventaja adivina que puede tener la otra?

Varianza del estadístico media muestral: $V(\bar{X})$. La dispersión de $\bar{X}$ es directamente proporcional a la dispersión de X e inversamente proporcional al tamaño n de la muestra.

Fórmula

$V(\bar{X}) = V(X)/n$

Nótese que la relación entre la variabilidad de las medias muestrales, $V(\bar{X})$, y el tamaño muestral, n, es inversa. Ello implica que cuanto mayor sea el tamaño n de la muestra, menos varían las medias muestrales $\bar{X}$. Ya intuíamos que a mayor tamaño de la muestra, mayor credibilidad de resultados, pero ahora una fórmula los relaciona.

Ejemplo 2.8: Supóngase que se están tomando muestras de la altura de niños. La variabilidad de las posibles medias muestrales será mayor si tomamos muestras de tamaño n=3 que si son de tamaño n=1000.

Lectura: Una flor no indica primavera.

Ejercicio 2.3

¿Es coherente esta situación?, ¿qué utilidad puede tener?

Contra-Ejemplo 2.9: No se tendrá más información si se mide 1000 veces al mismo niño. Para que una nueva observación aporte información completa deberá ser independiente de las observaciones previas.

Tampoco sorprende que a mayor variabilidad de la variable, mayor oscilación de la media $\bar{X}$ de una muestra a otra.

Ejemplo 2.10: En muestras de la altura de niños, la variabilidad de las medias será mayor si los niños tienen edades comprendidas de 5 a 15 años, que si todos tienen 8 años.

Ejemplo 2.11: Suponga que los ingresos de los titulados de una facultad aumentan con el tiempo que pasa desde que dejan la universidad. Si es así, la dispersión de X será mayor si se estudia el conjunto de todos los titulados, que si se estudia solamente los titulados en un cierto año. En consecuencia las medias muestrales, $\bar{X}$, obtenidas de la población total, fluctuarán más que las obtenidas de muestras de un solo curso.

2.5. Error típico

La varianza de $\bar{X}$ es el promedio de las distancias (cuadradas) con el centro: representa el error (cuadrado) esperado que se cometería en una muestra al estimar el parámetro poblacional a partir del valor muestral $\bar{X}$ obtenido.

Recuerde

La varianza de $\bar{X}$ proporciona el promedio de los errores al cuadrado.

Si la varianza de $\bar{X}$ informa del error "cuadrado", su raíz, dará el error típico de $\bar{X}$.

Fórmulas error típico de $\bar{X}$

A nivel teórico, conocida σ poblacional: $\sigma_{\bar{X}} = \sqrt{V(\bar{X})} = \frac{\sigma}{\sqrt{n}}$

A nivel práctico, a partir de la S muestral: $S_{\bar{X}} = \frac{S}{\sqrt{n}}$

Ejercicio 2.4

Suponga ahora que está interesado en conocer el promedio de hermanos de las familias de los alumnos del Ejercicio 2.2. Si esta muestra de n=5 fuera una muestra aleatoria (y, por tanto, representativa) de todas las familias, ¿qué error cabe esperar que tiene la media observada en la muestra al estimar la media poblacional? ¿Cómo lo interpreta?

Recuerde

El error típico de la media $\bar{X}$ es la desviación típica de la variable en estudio dividida por la raíz del número de casos.

Ejemplo 2.12: La siguiente frase "*los 100 niños tratados han tenido fiebre durante una media de 3 días; el error típico (o estándar) ha sido de 0.1 día*" hace inferencia hacia los valores de la población: se afirma que, al aplicar este tratamiento en todos los niños de la población origen, la media de duración de la fiebre es de 3 días y que el error esperado al decir que la media poblacional es de 3 días es de 0.1 día.

Debe quedar claro que se trata de un *error*, por tanto con connotación negativa. Nótese que mientras el término desviación típica no debería tener ninguna connotación, ni positiva ni negativa, ahora el error típico ya deja claro desde el primer momento que se trata de algo negativo, no deseable: el error que cabe esperar que se cometa al estimar el parámetro media poblacional a partir del estimador media muestral.

Ejemplo 2.13: La altura de las mujeres adultas tiene una distribución Normal de media μ=165cm, y desviación típica σ=7cm. Que la desviación típica sea de 7cm no es ni bueno ni malo, simplemente refleja una situación natural: para un ecólogo, será fuente de riqueza; para un fabricante de pantalones, un reto que superar. En cambio, si para estimar la altura media μ de las mujeres se calcula la media en una muestra de n= 100 mujeres, el error típico que conlleva la estimación es:

$$S_{\bar{X}} = \frac{S_X}{\sqrt{n}} = \frac{7}{\sqrt{100}} = 0.7\ cm$$

Este valor del error típico dice que la imprecisión de $\bar{X}$ al estimar μ es de 0.7cm.

Nota: de la misma manera que $\overline{X}$ estima μ, S estima σ y $S_{\overline{X}}$, $\sigma_{\overline{X}}$. Como en general no se conocerá σ, el error típico que se emplea es $S_{\overline{X}}$.

Ejercicio 2.5

Se estima en 4.4 puntos el incremento en la calidad de vida de la semana 0 a la 24 en 43 pacientes. Si la desviación típica observada ha sido de 1.2 puntos, ¿Cuánto vale el error típico? Interprete el resultado. Diga qué cuantifican, en este ejemplo, la desviación típica y el error típico.

El *error típico* habla del error esperado o promedio, ya que el error exacto que se comete en una muestra concreta permanece desconocido y puede ser más grande o más pequeño.

Nota: Formalmente no se puede interpretar el error típico como el promedio de los errores (es la raíz cuadrada del promedio de los errores cuadrados), pero a nivel práctico, decir que representa el error promedio o esperado es una buena aproximación.

A diferencia de la desviación típica, el error típico puede hacerse tan pequeño como se quiera: simplemente se trata de aumentar el tamaño de la muestra —siempre aleatoria.

El *error típico* habla de error aleatorio. Si Vd. dispone de una muestra aleatoria, aunque sea pequeña, sabrá cuantificar la oscilación originada por suerte, buena o mala.

Figura 2.3 El azar permite cuantificar la precisión del muestreo

2.6. ¿Desviación típica o error típico?

"¿Y qué debo utilizar, la desviación típica o el error típico?".

Esta pregunta no tiene razón de ser, ya que no son medidas alternativas para un mismo objetivo: la desviación típica es una medida descriptiva de cómo son los casos, mientras que el error típico es una medida del error asociado a un proceso inferencial. Así, se usa la desviación típica al describir los casos en los que se ha hecho el estudio (al inicio de "resultados"); y se usa el error típico al inferir (desde la muestra a la población) el efecto observado. Esta inferencia permitirá a otros científicos utilizar nuestros resultados.

Recuerde

La desviación típica es una medida de dispersión que describe los datos: ¿cómo son mis casos?

El error típico es una medida del error de estimación al hacer inferencia: ¿qué incertidumbre o ruido conlleva mi salto de la muestra a la población?

Lectura: La guía CONSORT aconseja la desviación típica para la tabla de descriptiva inicial y medidas de inferencia (como el error típico) para estimar el efecto de la intervención.

Recuerde

La desviación típica se usa al inicio, al describir la muestra; el error típico, al final, al inferir el resultado principal a la población objetivo.

La relación entre el tamaño muestral y el error típico no es inversamente proporcional. Sí que es inversa, pero hay una raíz por medio. Por tanto, si se quiere disminuir el error típico a la mitad, se deberá multiplicar por cuatro el tamaño muestral.

Ejemplo 2.14: El nivel de plaquetas en pacientes de una determinada enfermedad tiene una V(X)=2500 unidades^2. Si, para conocer su valor medio, se obtiene una muestra de 25 pacientes, el error típico del promedio es:

$$\sigma_{\bar{X}} = \sqrt{V(\bar{X})} = \sqrt{[V(X)/n]} = \sqrt{[2500\ u^2/25]} = \sqrt{100\ u^2} = 10\ u$$

En cambio, si se aumenta la muestra de 25 a 100 casos, el error típico es:

$$\sigma_{\bar{X}} = \sqrt{V(\bar{X})} = \sqrt{[V(X)/n]} = \sqrt{[2500\ u^2/100]} = \sqrt{25\ u^2} = 5\ u$$

Recoger cuatro veces más casos baja la oscilación a la mitad.

Recuerde

Si desea estimar un parámetro y dispone de un estimador insesgado, el error típico (SE: Standard Error) de este estimador, le informa del error esperado al afirmar que el valor del parámetro poblacional coincide con el valor del estimador obtenido en su estudio.

2.7. Estabilidad del conjunto

El hecho de que el error típico de la media $\overline{X}$ se vaya haciendo más pequeño a medida que aumenta el tamaño muestral indica una cierta estabilidad de los valores obtenidos en los grupos que se contrapone a la variabilidad de los correspondientes a las unidades.

Recuerde

La variabilidad de los individuos contrasta con la regularidad del conjunto.

Ejemplo 2.15: Pongamos que la probabilidad de nacer varón sea ½. El próximo nacimiento de Barcelona tiene esta probabilidad de ser varón. Pero no será mitad niño y mitad niña: o bien será niño o bien será niña. La incertidumbre es total. En cambio, podemos tener la tranquilidad de que el próximo año nacerán alrededor de un 50% de niños y un 50% de niñas en Cataluña. No le pediremos a un político que elabore un plan de contingencia por si, por puro azar, durante unos años sólo nacen bebés de uno de los dos géneros.

Ejemplo 2.16: A una persona en concreto o le toca la lotería o no le toca. La incertidumbre es absoluta. Pero el que organiza los juegos, además de contar con su tanto por ciento, puede tener la suerte de no repartir los premios gordos (o justo al revés). A medida que crece el número de apostantes, más estables son los resultados para el organizador.

2.8. Más propiedades de los estimadores *

De la misma forma que la esperanza de $\overline{X}$ sirvió para definir el concepto de estimador insesgado, se puede utilizar la oscilación de $\overline{X}$ para definir otras propiedades.

Definición

Un estimador es **convergente** si, a medida que crece el tamaño de la muestra, se acerca progresivamente al valor del parámetro.

Así, cuanto mayor sea la muestra, mejor será la estimación.

Ejemplo 2.17: $\overline{X}$ es un estimador convergente: V($\overline{X}$) disminuye a medida que aumenta n.

Definición

Entre dos estimadores insesgados, se dice que es **más eficiente** el que tiene menor error típico.

Analogía: Sean otros dos lanzadores con arco. Ambos insesgados. El de la izquierda tiene mayor dispersión alrededor de la diana, por lo que es menos eficiente (**Figura** 2.3).

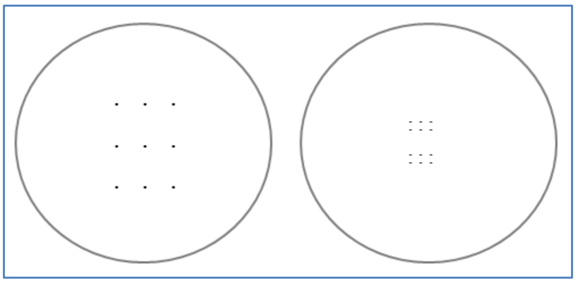

Figura 2.4 El arquero de la derecha es más eficiente porque su oscilación es menor.

Nota: Observe la connotación económica: el estimador más eficiente proporciona más información (tiene menos error aleatorio) para un mismo tamaño muestral (=coste). O también, permite obtener la misma cantidad de información con una muestra más pequeña (menor coste).

Ejemplo 2.18: Dos investigadores han diseñado dos experimentos para comparar la biodisponibilidad de dos preparaciones farmacéuticas alternativas. El primero ha obtenido dos muestras de voluntarios, administrando a cada una, una de las dos formulaciones. Luego compara las medias de las dos muestras ($\bar{X}_1 - \bar{X}_2$).

El segundo investigador ha recogido la información de las dos preparaciones en un único grupo, y calcula la media de las diferencias ($\bar{X}_D$), eliminando, de esta forma, la variabilidad debida al voluntario. Los gráficos muestran que, siendo ambos experimentos insesgados, el segundo es más eficiente.

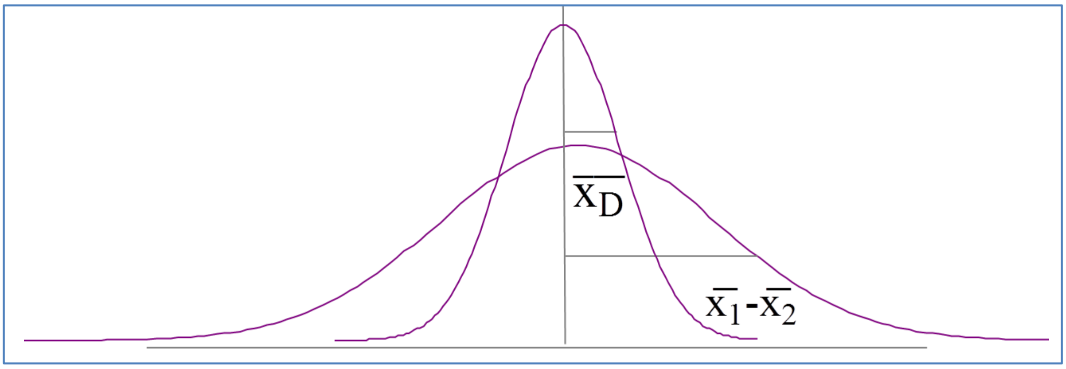

Figura 2.5 Ambos estimadores, $\boldsymbol{\bar{X}_1 - \bar{X}_2}$ y $\boldsymbol{\bar{X}_D}$, son insesgados pero $\boldsymbol{\bar{X}_D}$ es más eficiente

Lectura: La estadística permite cuantificar los errores aleatorios. Si le conviene que la Muestra Aleatoria Simple deje de ser "simple", ningún problema, un profesional de la estadística le ayudará a obtener el valor del error típico. Pero si la muestra deja de ser "aleatoria", debe discutir Vd. todos los posibles sesgos concebibles.

2.9. Estimación puntual

Definición

Al valor observado de un estimador en una muestra se le denomina **estimación puntual** del parámetro.

Ejemplo 2.19: En una muestra aleatoria de 9 personas, la presión arterial sistólica (PAS) ha tenido una media muestral $\bar{X}$ igual a 120 mmHg y una desviación típica muestral (S) de 12mmHg. Así, la estimación puntual de la PAS media en esta población ha sido de 120 mmHg. El error típico de esta estimación se puede cifrar en:

$$S_{\bar{X}} = S/\sqrt{n} = 12\, mmHg/\sqrt{9} = 4\, mmHg$$

Por lo tanto, hay una *señal* de 120 mmHg que está afectada por una *oscilación* de 4 mmHg.

Recuerde

El error típico informa del error esperado, pero el error exacto en una muestra concreta permanece desconocido, pudiendo ser inferior o superior.

Ejercicio 2.6

El descenso de la PAS tras la administración de un fármaco en una muestra de 16 pacientes ha tenido una media de 12 mmHg y una desviación típica de 8 mmHg. Calcule el error típico e interprete los resultados.

Ejercicio 2.7

Si hubiera deseado que el error típico hubiera sido de 1 mmHg, ¿Cuántos casos hubiera necesitado (desviación típica de 8 mmHg)?

2.10. Forma de la distribución del estadístico $\bar{X}$

Ya se ha dicho que las posibles medias muestrales $\bar{X}$ se distribuyen alrededor de la media poblacional μ=E(X) con una distancia promedio que cuantifica el error típico. Ahora bien, ¿qué forma tiene la distribución de $\bar{X}$?

Por las leyes de combinatoria y probabilidad, en general la muestra contendrá tanto valores superiores como inferiores a la media poblacional; y su media $\overline{X}$ será próxima a la media poblacional μ. También es posible obtener valores alejados, si bien será menos frecuente; de hecho, cuanto más se aleje $\overline{X}$ de μ, menos probable es observarla. La distribución Normal de Gauss-Laplace aplica a la $\overline{X}$ de una MAS.

Recuerde

La media muestral $\overline{X}$ se distribuye de acuerdo con la ley Normal Gauss-Laplace.

La distribución Normal aparece en variables que son el resultado de muchos factores o fuerzas que actúan independientemente y con influencias similares. Y eso es precisamente lo que es una media muestral $\overline{X}$, ya que cada observación de la muestra contribuye con el mismo peso o influencia. Queda por aclarar qué significa “muchos”: ¿cuántos casos se necesitan para que la distribución del promedio de una muestra se acerque a la ley Normal?

Nota técnica: El Teorema del Límite Central (TLC) establece que, si se toman muestras de tamaño n, de una población de media μ y desviación típica σ, a medida que crece n, la distribución de $\overline{X}$ se aproxima a la ley Normal con media μ y desviación típica $\sigma/\sqrt{n}$.

Ejercicio de navegación

Asegúrese de que su navegador soporta JAVA y observe en esta página cómo se comporta la media muestral $\bar{X}$ al crecer “n”.

Ejercicio 2.8

A partir de lo visto, ¿la distribución de la variable $\bar{X}$ cambia de forma cuando crece el tamaño de la muestra? ¿y la de X?

Ejercicio 2.9

¿Cómo cambia la forma de la distribución de la variable $\bar{X}$ cuando crece el tamaño de las muestras?

Ejemplo 2.20: La edad de los pacientes incluidos en un estudio sigue una distribución uniforme (aplanada, con el mismo número de casos en todas las franjas de edad). Si se toman muestras de tamaño n=30 y se calcula la media muestral $\bar{X}$ de la edad, la distribución de $\bar{X}$ se acercará a la Normal. La de la edad, sigue siendo la misma.

Ejemplo 2.21: La Presión Arterial Sistólica en los adultos sanos tiene una distribución que se asemeja bastante a la ley Normal. Si se toman muestras de tamaño n=3 y se calcula la media muestral $\bar{X}$ de la PAS, la distribución de esta media será Normal.

Recuerde

Las condiciones para poder creer que el promedio obtenido en una muestra $\bar{X}$ sigue una distribución Normal son, o bien muestra $n \geq 30$ o bien distribución Normal de la variable en estudio.

2.11. Intervalo 1-α de las medias muestrales $\bar{X}$

Se vio que la distribución Normal permite construir intervalos que contengan un determinado porcentaje de unidades o casos. Ahora, la variable en estudio es $\bar{X}$, por lo que, utilizando la Distribución Normal, se pueden construir intervalos que contengan un deseado porcentaje de las medias $\bar{X}$ que se podrían obtener en todas las posibles muestras.

Recuerde

Intervalo 1-α de $\bar{X}_n = \mu \pm z_{1-\alpha}\, \sigma/\sqrt{n}$

Intervalo 95% de $\bar{X}_n = \mu \pm 1.96\, \sigma/\sqrt{n}$

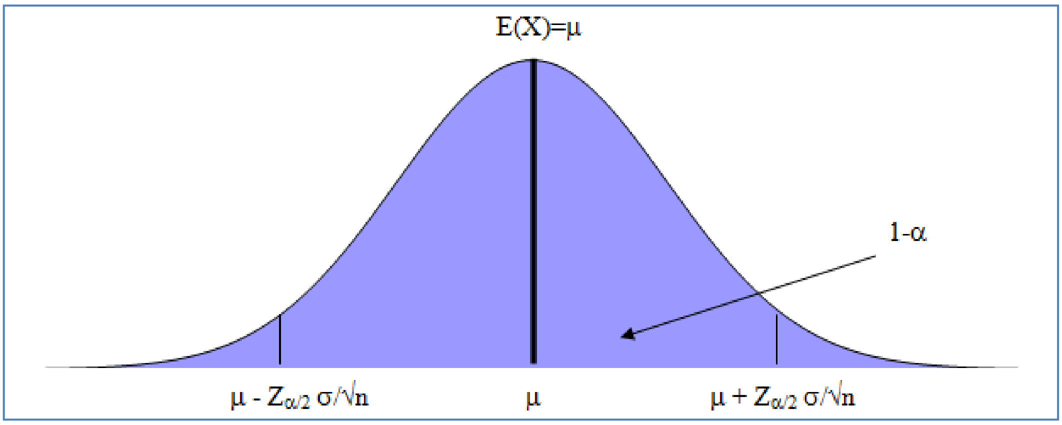

Figura 2.6 Distribución del estimador $\bar{X}$ alrededor del parámetro E(X) = μ. Una proporción 1 - α de las posibles medias muestrales $\bar{X}$ está incluida entre los límites indicados.

Para que este intervalo contuviera el 95% de las medias muestrales, el valor de la distribución Normal debía ser $Z_{\alpha/2} = Z_{0.025} = 1.96$.

Ejemplo de R

```
# Z0.025 (cola superior, lower.tail=FALSE)
> qnorm(p=0.025,lower.tail=FALSE)
[1] 1.959964
```

Ejemplo 2.22: La glucosa en sangre (X) sigue una distribución Normal de media μ=100 y desviación típica σ=10: $X \to N(100\ mg/ml\,, 10\ mg/ml)$

Se desean construir intervalos que contengan:

(i) el 95% de las unidades de la población;

(ii) el 95% de las posibles $\bar{X}$ de muestras de tamaño n=9; y

(iii) el 95% de las posibles $\bar{X}$ de muestras de tamaño n=100.

Los tres intervalos coinciden en que deben contener el 95% de sus unidades y dejar fuera el 5% (α=0'05). Pero se refieren a unidades totalmente diferentes, con distribuciones diferentes. En el primer intervalo las unidades son individuos; mientras que en el segundo y tercer ejemplo se trata de las medias muestrales que se obtendrían si se repitiera indefinidamente el proceso de tomar muestras de n=9 y n=100 de estos individuos.

Todas estas distribuciones (Figura 2.6) seguirán la ley Normal: al ser Normal la distribución de la glucosa en los casos, también lo es la distribución de la media $\bar{X}$, sea cual sea el número de casos. Todas tienen, también, la misma media. Pero cambia la dispersión: para el primer ejercicio, se trata de la desviación típica de la variable original, la glucosa en sangre, 10 mg/ml; mientras que para los dos restantes, se trata del error típico, debiéndose dividir la desviación típica por la raíz del número de casos respectivos:

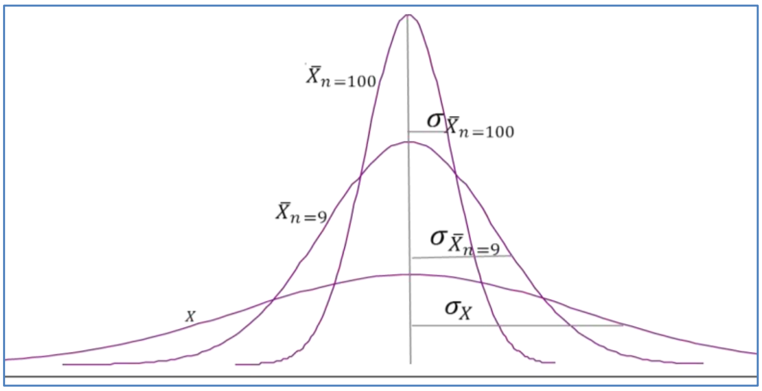

Figura 2.7 Distribución de las variables X, $\bar{X}_{n=9}$ y $\bar{X}_{n=100}$

(i) $X \to N(100\ mg/ml\,, 10\ mg/ml)$

(ii) $\bar{X}_{n=9} \to N\left(100, 10/\sqrt{n}\right) = N\left(100, 10/\sqrt{9}\right) = N(100\ mg/ml\,, 3.33\ mg/ml)$

(iii) $\bar{X}_{n=100} \rightarrow N\left(100, 10/\sqrt{n}\right) = N\left(100, 10/\sqrt{100}\right) = N(100\ mg/ml\,, 1\ mg/ml)$

Los límites de los intervalos se pueden calcular utilizando la Normal: $Z_{0.025} = 1.96$

i) Intervalo que contiene el 95% de las glicemias individuales, X:

$$\mu \pm z_{\alpha/2}\,\sigma/\sqrt{n} = 100 \pm 1.96 \cdot 10 = 100 \pm 19.6 = [80.4, 119.6]$$

ii) Intervalo del 95% de las medias ($\bar{X}_{n=9}$) de las muestras de n=9 individuos, $\bar{X}_{n=9}$:

$$\mu \pm z_{\alpha/2}\,\sigma/\sqrt{n} = 100 \pm 1.96 \cdot 10/3 = 100 \pm 6.53 = [93.47, 106.53]$$

iii) Intervalo del 95% de las medias ($\bar{X}_{n=100}$) de las muestras de n=100 individuos, $\bar{X}_{n=100}$:

$$\mu \pm z_{\alpha/2}\,\sigma/\sqrt{n} = 100 \pm 1.96 \cdot 10/10 = 100 \pm 1.96 = [98.04, 101.96]$$

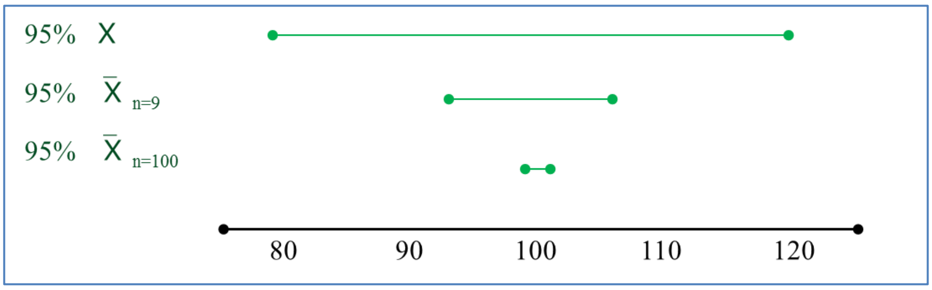

Figura 2.8 $IC_{95\%}$ de X, $\bar{X}_{n=9}$ y $\bar{X}_{n=100}$

Nota: Observe, una vez más, la mayor variabilidad de las muestras de menor tamaño.

Este ejemplo muestra cómo obtener, a partir de los valores poblacionales de media [E(X)=μ] y varianza [V(X)], dónde estarán los valores de la media muestral ($\bar{X}$) en el 95% de las posibles muestras. Puede ser interesante, pero tiene poca utilidad práctica, ya que el problema habitual es justo al revés: conocidos los estimadores muestrales de media ($\bar{X}$) y varianza (S^2), ¿qué se sabe de la media poblacional E(X) = μ? Esta interesante pregunta se contesta en el siguiente tema.

Ejercicio 2.10

El cociente de inteligencia (CI) sigue en terminos generales, una N(100,15). Si se recolectaran muchas muestras de tamaño n=9 y en cada muestra j se calculara su media $\bar{X}_j$:

a) ¿Cómo variarían las medias $\bar{X}_j$ de la muestras?

b) En una facultad de Medicina, se ha recogido una muestra de tamaño n=9 y se ha observado $\bar{X}_j$=104, ¿se trata de (1) un ejemplo aceptablemente típico; o (2) especialmente afortunado, muy cerca de μ; o (3) tan raro y alejado que se sospecha que estos alumnos no son de aquella población?

c) Repetir los dos apartados anteriores, pero con n=25 y n =225.

Soluciones a los ejercicios

1.1. Sobre ambos. Escogeremos sus indicadores según la forma de la distribución, pero siempre debemos informar de ambos. Usualmente, la comunidad científica enfoca su interés en los valores centrales para poder resumir o representar a los casos. Pero hay que hacer el esfuerzo de cuantificar también el grado de dispersión. E incluso conocer la forma de la distribución.

1.2. Cualquier ejemplo es válido. También sería terriblemente aburrido “decir toda la verdad” sobre la carga viral. Nótese, en cambio, que no lo sería sobre el género (“53 fueron del género masculino y 47 del femenino”). Y quizás tampoco sobre el número de infecciones oportunistas (“2523 casos no presentaron ninguna; 48 tuvieron una; 7, dos y 1 caso, tres infecciones”).

1.3. En el estudio de los odontólogos, la unidad es “visita a la consulta”, mientras que en la población general, la unidad es “habitante”. Como hay habitantes que van al dentista más veces que otros, éstos estarán sobre-representados en un estudio en el que se seleccionen “visitas”. Nótese que aquellos que nunca van al dentista tienen una probabilidad nula de ser seleccionados. En resumen, porque hablan de unidades diferentes, el estudio de los dentistas (“visitas”) ofrecerá cifras distintas (en este ejemplo, más altas) que el de la población general (“ciudadanos”).

1.4. Los estudios de satisfacción hospitalaria se basan en las altas hospitalarias, mientras que el estudio del defensor se basó en ciudadanos. Igual que en el ejercicio anterior, aquellos ciudadanos que van menos a los centros públicos, tienen una probabilidad menor de ser seleccionados. En resumen, cabe esperar mayor satisfacción entre las “altas” hospitalarias, que entre los habitantes.

1.5. La unidad del primer estudio son los “pacientes ingresados”, pero la del segundo las “camas ocupadas”. El primero informará sobre la frecuencia de infección nosocomial en un paciente que acuda a ese hospital; pero el segundo, de la probabilidad de que cierto paciente ingresado cierto día y ocupando una cama, tenga dicha condición. Igual que antes: si los pacientes que desarrollan infecciones nosocomiales permanecen más tiempo en el centro y la selección se hace a partir de las “camas ocupadas”, las cifras de infección serán mayores.

1.6. Por ejemplo, en un estudio del perfil lipídico en 41 pacientes con HIV tratados, la media del colesterol total fue de 4.51 mmol/l. Se trata de la media obtenida en la muestra, y que informa sobre el valor de la esperanza (media en la población), que es desconocida.

1.7. La muestra queda configurada por los casos concretos seleccionados; la población muestreada es aquella de la cual se hubiera obtenido por meros mecanismos aleatorios, la muestra; la población objetivo, todos los fumadores de esa región sanitaria; y la población externa, los fumadores de la zona dónde se quieren aplicar los resultados.

1.8. No, si el objetivo es utilizar los resultados en esos centros. Pero si desea utilizarlos en otros centros con diferente gravedad habría sesgo.

1.9. Se trata del sesgo de publicación: al publicar sólo lo que ha resultado significativo, se da menor oportunidad a difundir resultados sobre no-eficacia (volveremos a este tema).

2.1. Si se desea utilizar la media muestral (X) para conocer la media poblacional ($E(X)=\mu$) es bueno que la distribución de $\bar{X}$ se disponga alrededor del auténtico valor de μ. Dicho al revés: sería peor que se distribuyera alrededor de cualquier otro valor. Y cuanto más alejado de μ estuviera, peor.

2.2. Media $\bar{X} = \sum x_i/n = (1+2+3+4+5)/5 = 3$

Varianza $S^2 = \sum(x_i - \bar{X})^2/(n-1) = [(1-3)^2 + (2-3)^2 + (3-3)^2 + (4-3)^2 + (5-3)^2]/4 = 10/4 = 2.5$

$S^2 = \left[\sum x_i^2 - (\sum x_i)^2/n\right]/(n-1) = [(1^2 + 2^2 + 3^2 + 4^2 + 5^2) - (1+2+3+4+5)^2/5]/4 = [55 - 225/4] = 10/4 = 2.5$

$S = \sqrt{S^2} = \sqrt{2.5} \approx 1.58$

Es decir, la media muestral es 3 hermanos; la varianza muestral, 2.5 hermanos² y la desviación tipo muestral es de aproximadamente 1.6 hermanos. Podemos imaginar que la distancia (o desvío) de una familia "típica" con la media es de 1.6 hermanos.

La primera es más intuitiva. La segunda (esto no tiene porqué saberlo) es computacionalmente más eficiente y más exacta.

2.3. Es coherente: cuantos más casos se tiene, de más información se dispone: hay menos error aleatorio. Es útil (y, por tanto, deseable) en el sentido de que un mayor esfuerzo en la recolección de datos se ve recompensado por menor oscilación de las estimaciones.

2.4. $S_{\bar{X}}^2 = S^2/n = 2.5/5 = 0.5$; $S_{\bar{X}} = \sqrt{S_{\bar{X}}^2} = \sqrt{0.5} \approx 0.707$

Si se afirma que la media de la población es de 3 hermanos (es decir, si decidimos aproximarnos a la media poblacional a partir de la media muestral), el error esperado al hacer esta afirmación es de 0.7 hermanos.

2.5. $S_{\bar{X}} = \sqrt{S_{\bar{X}}^2} = S/\sqrt{n} = 1.2/\sqrt{43} = 0.18$

La señal obtenida ha sido 4.4 y el error asociado que lleva esta señal es de 0.18.

La desviación típica dice que el incremento en calidad de vida observado en esta muestra tiene una distancia promedio de todos los casos de 1.2 al centro, estimado en 4.4. El error típico en cambio, habla de la oscilación esperada del estimador de la media; es decir, del error esperado al decir que la media muestral es igual que la poblacional.

2.6. $S_{\bar{X}} = \sqrt{S_{\bar{X}}^2} = S/\sqrt{n} = 8/\sqrt{16} = 2$

La señal obtenida ha sido de 12 mmHg y el error asociado que lleva esta señal es de 2 mmHg.

2.7. Si desea que $S_{\bar{X}} = 1 \rightarrow 8/\sqrt{n} = 1 \rightarrow n = 8^2 = 64$. Una vez más, si desea reducir a la mitad la oscilación del estimador, debe multiplicar por 4 el tamaño muestral (64=16·4).

2.8. A medida que crece el tamaño muestral, lo que va cambiando de forma es la distribución de la variable media muestral $\bar{X}$. La distribución de los valores observados, es decir, lo que se llama la distribución de la variable en estudio es siempre la misma para todos los casos, haya 3, 50 ó 1000. Si no tiene claro que lo que cambia es la distribución de la media muestral $\bar{X}$, no la de X, repita la última navegación.

2.9. Su centro, la esperanza, no cambia, pero sí que lo hace la dispersión y puede hacerlo la forma. La dispersión, cuantificada por el error típico, se va haciendo más pequeña a medida que crece el tamaño muestral (la reducción es proporcional al incremento de $\sqrt{n}$). La forma, en el caso de variables que no siguen una distribución Normal, se aproxima cada vez más a la de esta distribución (en el caso de variables que siguen la ley Normal, ya tiene esta distribución para cualquier n).

2.10. a) $V(\bar{X}) = V(X)/n = 15^2/9 = 25\ u^2 \rightarrow \sigma_{\bar{X}} = 5\ u$

La variabilidad de las medias muestrales es la tercera parte de la variabilidad de la variable.

b) $[\bar{X}_i - E(X)] = [104 - 100] = 4\ u$; cifra "razonable" ya que su valor esperado era 5 u. Por tanto diríamos que la opción correcta es la "(1) un ejemplo aceptablemente típico".

c) Si n = 25, $V(\bar{X}) = V(X)/n = 15^2/25 = 9\ u^2 \rightarrow \sigma_{\bar{X}} = 3\ u$

$[\bar{X}_i - E(X)] = [104 - 100] = 4\ u$; cifra "razonable", ya que su valor esperado es 3 u.

Si n = 225, $V(\bar{X}) = V(X)/n = 15^2/225 = 1^2\ u^2 \rightarrow \sigma_{\bar{X}} = 1\ u$

$[\bar{X}_i - E(X)] = [104 - 100] = 4\ u$, NO es una cifra "razonable", ya que su valor esperado es 1u. Por tanto, diríamos que la opción correcta es la "(3) tan raro y alejado que se sospecha que estos alumnos no son de aquella población".